Supermicro ICR Recipe
For 1U Twin™ 'Department Cluster'
with ClusterVision ClusterVisionOS™

Version 1.4
6/25/2009

Table of Contents

1. System Configuration

Bill Of Materials (Hardware) The primary bill of materials, orderable from Supermicro's distributors, is shown in Table 1.   The BOM in Table 1 corresponds to that used for the cluster certification.

Quantity	Item	Manufacturer	Model
16	SuperServer 1U Twin™ with QDR InfiniBand	Supermicro	SS6016TT-IBQF
2	24-port Layer 3 1/10-Gigabit Ethernet Switch	Supermicro	SSE-G24_TG4
1	Gigabit Ethernet Switch Stacking Cable	Supermicro	MCC4L30-001
2	Gigabit Ethernet CX4 Stacking Module	Supermicro	AOM-SSE-X2C
1	36-port InfiniBand Switch	Mellanox	MTS3600
32	InfiniBand Cable, 1m, QSFP	Mellanox	MCC4Q30C-001
64	Intel® Xeon® Processor	Intel Corp.	Intel® Xeon® Processor X5550 (Nehalem)
192	2GB DDR-3-1066 ECC Registered Memory	Hynix	HMT125R7AFP4C-G7TB
32	3.5" SATA 250GB Hard Disk Drive	Seagate	Barracuda ES.2 ST3250310NS

Table 1: 1U Twin™ Cluster- Bill of Hardware Materials

This recipe provides a template that resellers and end-users can customize for their application specific needs. In addition to the hardware items in Table 1, the following are either required or useful during the installation process: USB keyboard and mouse, USB DVD-ROM drive, and USB 4 port hub. Please refer to the SS6016TT-IBQF User Manual for instructions on assembling the unit. Although the BOM specifies the 1UTwin™ system, the 2U Twin²™ uses the same server motherboard, BIOS and Firmware and thus qualifies as 'materially identical' under the rules of the ICR program. Therefore, both 1U Twin™ and 2U Twin²™ based clusters using the X8DTT-IBQF and X8DTT-IBXF motherboards are ICR certified with ClusterVision ClusterVisionOS™.

Bill Of Materials (Software) The software bill of materials consists of a single DVD from ClusterVision, ClusterVisionOS™ 4.0 (see Table 2 below). In order to build and deploy ICR certified clusters with ClusterVisionOS™, a software and support license must be purchased from ClusterVision. This can be obtained in multiple ways as described at http://www.clustervision.com/products_clustervisionos.php .

Distributed By	Description	File Name or Location
ClusterVision	ClusterVisionOS 4.0	ClusterVisionOS 4.0 DVD

Table 2: 1U Twin™ Cluster- Bill of Software Materials

Bill Of Materials (Intel Cluster Ready License) The system vendor is required to run the Intel Cluster Checker tool both before it leaves the factory and after installation at the end user site.   The Cluster Checker tool requires a license file to be installed on the cluster. The license file can be obtained free of charge directly from Intel at http://www.intel.com/go/cluster (registration required) if desired. Alternatively, the reseller can use Supermicro's pass through license included in the download bundle from the Supermicro web site along with this recipe.

Bill Of Materials (Download Bundle) The system vendor is required to run the Intel Cluster Checker tool both before it leaves the factory and after installation at the end user site (see Section 4 below). The Cluster Checker tool requires several files which are bundled together as a single download from the Supermicro web site along with this recipe. The bundle includes: fingerprint files, XML Config & Output files, Supermicro ICR license file (COM_*), the cluster certification certificate, a copy of this recipe document and a README file.

2. Firmware and BIOS Settings

Once the hardware and software system components have been obtained and the servers have been assembled and racked, you can begin the system configuration. Refer to the 1U Twin™ (SS6016TT-IBQF) System User Manual for details on the server assembly procedure.

Connect a keyboard, mouse and monitor to each server in turn and configure the BIOS as follows.   Enter the BIOS setup as described in the system User Manual.   First ensure that the BIOS rev is 1.0b, or dated 4/20/09 or later.   If not, download the latest BIOS from the Supermicro web site and update the BIOS as described on the BIOS download web page. Next,

go to the far right tab on the main menu and enter 'Load optimized defaults'. Then set each BIOS parameter as follows, where '/' indicates a submenu in the BIOS setup.

- Advanced / Processor&ClockOptions / SimultaneousMultithreading = Disabled
- Advanced / Processor&ClockOptions / Intel EIST Technology = Disabled
- Advanced / IDE-SATA-FloppyOptions / ConfigureSATA = AHCI
- Advanced / PCI-PnPConfiguration / Load Onboard LAN 1 Option ROM = Enabled
- Advanced / ACPIConfiguration / ACPIVersionFeatures = ACPI2.0
- Advanced / IPMIConfiguration / LAN Configuration / IPAddress / IPAddrSource = Static
- Boot / BootDevicePriority / 1stBootDevice = Network: IBA GE Slot

On the head node only:

- Advanced / PCI-PnPConfiguration / Load Onboard LAN 1 Option ROM = Disabled
- Boot / BootDevicePriority / 1stBootDevice = HDD

Power off each server after the BIOS update is complete. The next step is to install ClusterVisionOS™ 4.0 on the cluster head node.

3. Software Installation

Please check the following chapters in the ClusterVisionOS™ 4.0 Administrator Manual:

Chapter 2: Installing ClusterVisionOS™
Chapter 3: Cluster Management with ClusterVisionOS™
Chaper 4: Configuring Your Cluster
Chapter 5: Power Management
Chapter 6: Node Provisioning

4. Verify a Correct Cluster Build

Responsibilities of Reseller

Under the terms of the pass through certification clause of the ICR program agreement (between Supermicro and Intel), the reseller shall run the cluster checker tool against the fingerprint files & XML configuration file provided by Supermicro along with this recipe. The cluster must pass the tests listed below before it leaves the reseller's factory. In addition, the reseller must make provision to re-run the tests once the cluster is installed at the end user site. This helps ensure that the system is functional after shipping (loose cables for example). The cluster checker tests are not burdensome. They could typically be incorporated into the reseller's system burn in procedure for example.

Cluster Validation with ICR

You are now ready to validate extended cluster functionality with the ICR Cluster Checker. Refer to the ICR documentation and Cluster Checker manual for details on the procedures outlined below. The ICR documents can be found at http://www.intel.com/go/cluster. Begin the cluster validation process as follows:

To make a freshly installed CVOS4 cluster Intel Cluster Compliant, perform the following steps:

Contents

    * 1 Install extra packages on master and image
    * 3 Infiniband
    * 4 Node information
    * 5 User-mode Checks
    * 6 Root-mode Checks

Install extra packages on master and image

 yum install cvos-config-intelcompliance-master
 yum --installroot=/cvos/images/default-image install cvos-config-intelcompliance-slave

Repeat the last command for all software images

Infiniband

If the setup includes an Infiniband fabrics, make sure OFED tools have been properly installed and configured. Note that the OpenSM daemon must not be running on a compute node!

As of Intel Checker 1.3, the OpenIB test is optional and therefore not performed by default in compliance/certification mode. To force OpenIB check, uncomment the line <include_module>openib</include_module> in /home/cvsupport/clk_results/recipe-user.xml and adapt the test parameters to your current setup. You will probably need to customize the adapter device name and the port status in the adapter container of the openib test.

Node information

Before running the cluster check tool, be sure the host list in /home/cvsupport/clk_results/nodelist fits your hardware setup. Also check that cvsupport user home directory is mounted on all hosts, otherwise mount it manually by issuing the command mount -a.

User-mode Checks

To perform the check, as the cvsupport user, do the following:

 module initadd intel-cluster-checker intel-cluster-runtime
 module load intel-cluster-checker intel-cluster-runtime
Then execute the instruction below to check for certification in non-privileged user mode:

 cluster-check --certification 1.1 /home/cvsupport/clk_results/recipe-user.xml

You may have to tweak some values for performance tests. They are mostly bandwidths and latencies for tests like memory_bandwidth_stream, imb_pingpong_intel_mpi, intel_mpi_testsuite, hpcc, etc. This could also be other units like FLOPS for mflops_intel_mkl.

Root-mode Checks

Then as root add Intel Checker and Runtime to your environment,

 module load shared intel-cluster-checker intel-cluster-runtime

And run the following command to generate files containing a list of all installed packages on the head node and one of the compute nodes (default is the first node in the node list file):

 cluster-check --packages /home/cvsupport/clk_results/recipe-root.xml

This script will produce two output files (with a timestamp). Copy the ouput files to /home/cvsupport/clk_results/{head-packages.list,node-packages.list} respectively.

Then run (as root):

 cluster-check --certification 1.1 /home/cvsupport/clk_results/recipe-root.xml

You may run here into hardware difference issues, especially in the dmidecode check. You can fix this by creating group(s). Lets assume that one of the compute nodes (node003) has a different memory device than the other nodes. If run with the default configuration, the checker will fail on dmidecode. To solve the problem, create a group for that node (add '# group: differentmemdev' after node003) and refer to that in dmidecode using the <group> tag (<group name="differentmemdev"/> in the dmidecode part of /home/cvsupport/clk_results/recipe-root.xml).

In order to get checker-level ICR certification, both runs should return "Check has Succeeded"

Power and Cooling Requirements The cluster configuration described in the BOM was measured as consuming 11 kW at 208V AC running Linpack on each node. This is expected to be highly configuration dependant, for example depending on the CPU power rating, number of DIMMs installed and so on. Therefore this measurement should be used only as a reference point.

Permissible Configuration Variants The pass through cluster certification is valid for certain variations on the configuration detailed in this recipe. Different processor types, memory manufacturer, density and number of DIMMs, disk drive manufacturer, capacity and count are permitted. Different models of server motherboard and Rocks software stack are not permitted. In that case the reseller may complete the full certification procedure themselves and apply to the ICR program web site for certification certificate. Important minimum configurations limits include

- Memory – 1 GB per processor core
- Disk – the head node must have 65GB of available storage. Available storage is formatted capacity minus swap.

5. Contacts

The reseller is responsible for first level product support with the end user. A comprehensive 2nd / 3rd level support package (for the reseller) is available as follows:

- Initial support requests can be made through the Supermicro support center via phone, email or web. It is essential to have the unit serial number on hand in order to process the request. The Supermicro support center can be found at http://www.supermicro.com. Resellers may use the SuperServer support hotline.
- The Supermicro support center will make a determination as to whether the problem is hardware related or software related. Hardware related issues will be handled by the Supermicro support center in the usual fashion.
- Software related issues will be redirected to ClusterVision. It is essential to have purchased ClusterVisionOS™ from ClusterVision to be licensed to use the software and receive support.

6. Release Notes

- When creating user accounts, in order for the auto-mounter to work correctly, the home directory must be specified as /export/home/*. After the user account is created on the head node, issue the command 'rocks sync users' to replicate across the compute nodes.